REVSTAT – Statistical Journal
Volume 9, Number 3, November 2011, 199–212

THE GARMAN–KLASS VOLATILITY ESTIMATOR REVISITED

Author: Isaac Meilijson
– School of Mathematical Sciences,
Raymond and Beverly Sackler Faculty of Exact Sciences,
Tel-Aviv University, 69978 Tel-Aviv, Israel
meilijson@math.tau.ac.il

Received: December 2010 Revised: May 2011 Accepted: June 2011

Abstract:

- The Garman–Klass unbiased estimator of the variance per unit time of zero-drift Brownian Motion, is quadratic in the range-based financial-type data $CLOSE-OPEN$, $MAX-OPEN$, $OPEN-MIN$ reported on regular time windows. Its variance, 7.4 times smaller than that of the common estimator $(CLOSE-OPEN)^2$, is widely believed to be the minimal possible variance of unbiased estimators. The current report disproves this belief by exhibiting an unbiased estimator in which 7.4 becomes 7.7322. The essence of the improvement lies in data compression to a more stringent sufficient statistic. The Maximum Likelihood Estimator, known to be more efficient, attains asymptotically the Cramér–Rao upper bound 8.471, unattainable by unbiased estimators because the distribution is not of exponential type.

  Beyond Brownian Motion, regression-fitted (mean-1) quadratic functions of the more stringent statistic increasingly out-perform those of $CLOSE-OPEN$, $MAX-OPEN$, $OPEN-MIN$ when applied to random walks with heavier-tail distributed increments.

Key-Words:

- *Garman–Klass; Brownian Motion; volatility; estimation.*

AMS Subject Classification:

- 62F10, 62P05.

1. INTRODUCTION

Consider a mean-zero Brownian Motion with constant unknown unit-time variance σ^2, monitored over disjoint regular intervals of time for each of which the initial (*OPEN*), final (*CLOSE*), maximal (*MAX*) and minimal (*MIN*) values are reported. The Garman–Klass [5] variance estimator, introduced three decades ago, achieves the accuracy in estimating σ^2 that the classical, natural estimator *average* $(CLOSE-OPEN)^2$ does in 7.4 times the observation period. This unbiased variance estimator is the minimum-variance unbiased quadratic function of the spreads $c = CLOSE-OPEN$, $h = MAX-OPEN$, $l = MIN-OPEN$ (for *close, high, low*). As will be shown, range data $S_1 = (c, h, l)$ can be compressed further without loss of sufficiency, yielding an unbiased variance estimator with efficiency 7.73 with respect to c^2. There is not much room for further improvement, as the Cramér–Rao bound makes 8.5 out of reach. Rogers & Satchell [9] suggested another unbiased estimator of σ^2, with efficiency 6 with respect to c^2, that is unbiased even for general unknown drift. We do not attempt here to compress range data for non-zero drift.

As stressed repeatedly, volatilities change over time and past data should be given decaying importance, as in GARCH-type estimators. The present paper deals with constant volatility only, emphasizing efficiency as a means of making do with short observed histories.

A coarser (but incomplete) sufficient statistic. Consider the triple $S_2 = (C, H, L)$ where $C = |c|$, $(H, L) = (h, l)$ if $c > 0$, while $(H, L) = -(l, h)$ if $c < 0$. Without loss of relevant information about the variance, the Brownian Motion trajectory $\{B(t);\ t \in (0,1)\}$ may be replaced by the flipped path $\{W(t);\ t \in (0,1)\}$, defined as $W(t) = B(0) + [B(t) - B(0)]\,\mathrm{sign}(B(1) - B(0))$. That is, the three interval lengths $(-L, C, H-C)$, in fact the further compression $\big(C,\ \min(-L, H-C),\ \max(-L, H-C)\big)$, determined by (c, h, l), carry all relevant information contained in (c, h, l) about σ^2, but *do not determine* (c, h, l). Although intuitively clear after some thought, sufficiency of $\big(C, \min(-L, H-C), \max(-L, H-C)\big)$ can be formally inferred from Siegmund's [11] representation displayed as (A.1) in the sequel. The Rao–Blackwell theorem [3, 8] claims that under these conditions, for every S_1-based unbiased estimator of some arbitrary parameter there is an S_2-based unbiased estimator with smaller variance — strictly smaller unless the two coincide. As will be seen, the Garman–Klass estimator is a function of S_2, so the Rao–Blackwell improvement leaves it invariant. However, the Garman–Klass estimator, best among the quadratic function of S_1, is not best possible as a function of S_2. Had S_2 been a complete minimal sufficient statistic, Garman–Klass and the proposed estimator would have equally been the UMVUE (uniformly minimum variance unbiased estimator) of the parameter. However, C^2 and $2\,[(H-C)^2 + L^2]$ are different unbiased estimators of σ^2. Hence, S_2

(whether minimal sufficient or not) is not complete. Loose some, win some: we will only conjecture rather than claim optimality of the proposed S_2-based quadratic unbiased estimator of σ^2; on the other hand, the exchangeability property under which $(-L, C, H-C)$ and $(H-C, C, -L)$ are identically distributed, justifies searching for the best quadratic function of $(-L, C, H-C)$ among those that are linear combinations of four rather than six quadratic terms.

Four basic quadratic unbiased variance estimators. Consider

$$\hat\sigma_1^2 = 2\big[(H-C)^2 + L^2\big]\,, \qquad \hat\sigma_2^2 = C^2\,, \qquad \hat\sigma_3^2 = 2(H-C-L)C\,, \qquad \hat\sigma_4^2 = -\frac{(H-C)\,L}{2\log(2)-\frac{5}{4}}\,. \tag{1.1}$$

The rationale for the somewhat bizarre coefficients is that each of these four terms is an unbiased estimator of σ^2, with respective variances

$$\mathrm{Var}(\hat\sigma_1^2) = 0.797943\,\sigma^4\,, \qquad \mathrm{Var}(\hat\sigma_2^2) = 2\,\sigma^4\,, \qquad \mathrm{Var}(\hat\sigma_3^2) = 0.504753\,\sigma^4\,, \qquad \mathrm{Var}(\hat\sigma_4^2) = 1.004876\,\sigma^4\,. \tag{1.2}$$

The proposed variance estimator vis à vis Garman–Klass. The proposed estimator $\hat\sigma^2 = \sum_1^4 \alpha_i\,\hat\sigma_i^2$ assigns to these four terms respective weights

$$\alpha_1 = 0.273520\,, \quad \alpha_2 = 0.160358\,, \quad \alpha_3 = 0.365212\,, \quad \alpha_4 = 0.200910\,, \tag{1.3}$$

and achieves variance $\mathrm{Var}(\hat\sigma^2) = 0.258658\,\sigma^4$. The Garman–Klass estimator

$$\hat\sigma_{GK}^2 = 0.511(h-l)^2 - 0.019\big(c(h+l) - 2\,h\,l\big) - 0.383\,c^2 \tag{1.4}$$

happens to pool these four basic estimators too, so the Rao–Blackwell theorem does not rule out the possibility that it coincides with $\hat\sigma^2$. However, as argued earlier, the two do not agree, and $\hat\sigma_{GK}^2 = \sum_1^4 \beta_i\,\hat\sigma_i^2$ pays a price for being quadratic in (c, h, l). Its coefficients are given by

$$\begin{aligned}
\beta_1 &= \frac{0.511}{2} = 0.2555\,,\\
\beta_2 &= 0.511 - 0.383 - 0.019 = 0.1090\,,\\
\beta_3 &= 0.511 - \frac{0.019}{2} = 0.5015\,,\\
\beta_4 &= 2\,\big(0.511 - 0.019\big)\left(2\log(2) - \frac{5}{4}\right) = 0.1340\,,
\end{aligned} \tag{1.5}$$

that achieve $\mathrm{Var}(\hat\sigma_{GK}^2) = 0.27\,\sigma^4$.

Maximum Likelihood variance estimators and Fisher information. In principle, giving up on the requirement of unbiasedness, the computer-intensive maximum likelihood estimator (MLE) of σ^2 by Magdon-Ismail & Atiya [7] could have been a competitor, since MLE's are functions of any sufficient statistic.

However, this estimator is based on (h, l) rather than on (c, h, l). Magdon-Ismail & Atiya report that their estimator has variance slightly higher than Garman–Klass'. Variance estimators other than Garman–Klass and Rogers–Satchell have been suggested in the literature, some for unknown drift, range-based (based on *MAX* and *MIN* but not on *OPEN* and *CLOSE*, e.g., Alizadeh, Brandt & Diebold [1], Christensen, Podolskij & Vetter [4]) or otherwise (e.g., noisy or lattice measurements), but not unbiased — the subject matter of this paper. There is no theoretical limit as to how accurately can σ^2 be estimated, as its value is a.s. deterministically imprinted into the trajectory of B on any time interval of positive length.

The joint generating function of (c, h, l) is presented by Garman & Klass as an infinite series, from which these authors derived all pertinent second and fourth degree moments.

Ball & Torous [2] developed an infinite-series formula for the joint density of (c, h, l) and used it to construct numerically the MLE of σ^2. They report estimated efficiency of the MLE for a selection of sample sizes, basing each value on a simulation sample size of 1000 runs, a great achievement in 1984, but insufficient for delicate comparisons. The Fisher information was numerically re-evaluated via the formula by Siegmund quoted earlier, exhibited as (A.1) in the sequel. The inverse of the Fisher information is the Cramér–Rao lower bound for the variance per time-window of any unbiased estimator of σ^2, for any sample size. It is also the asymptotic variance of the (not necessarily unbiased) MLE of σ^2. Its value turns out to be 0.2361. This is the benchmark with which C^2's 2, Garman–Klass' 0.27 and the proposed estimate's 0.258658 variances should be compared.

For our problem, the Cramér–Rao bound 0.2361 is not attained by unbiased variance estimators: disproving exponentiality of a family of distributions. Under proper regularity assumptions (see Joshi [6]), the Cramér–Rao bound is attained if and only if there is a linear relationship between the estimator and the score function (derivative with respect to the parameter of the logarithm of the density). However, for this to happen, there must exist a linear relationship between the score functions evaluated at different values of the parameter. It was ascertained numerically that this is not the case. In other words, the model is not of exponential type. We don't know whether the sufficient statistic S_2, shown above not to be complete, is minimal sufficient. As a result of all of these considerations, the proposed estimator may not be of minimal variance.

Since both the proposed and Garman–Klass' estimators are averages over time-windows, their variances per time-window are independent of sample size. It is conceivable, and Ball & Torous have provided evidence in this direction, that the MLE has variance per time-window that decreases as the sample size increases, so for small sample sizes the proposed estimator has in practice no competitor.

Moreover, since the BM model doesn't really hold in practice, a broader contribution of this paper is the introduction of more efficient quadratic statistics on which to base practical estimators. Simulation results for random walks with t-distributed increments are reported in Section 3.

2. DERIVATION

Following the steps of Garman & Klass, all second and fourth order moments of (C, L, H) will be identified. Some of these will be quoted from Garman & Klass, some will be derived once the joint densities of (C, H) and (C, L) are explicitly presented, and some will require some additional argument. Although it would perhaps be more natural to work only with the exchangeable variables $\Delta = H - C$ and $\delta = -L$, work will be performed on the variables H and L as well, in order to link more easily with Garman & Klass' triple (c, h, l).

2.1. The joint densities of C and each of H and L: four unbiased estimators

Assume throughout the computations that the drift is 0 and the variance per unit time is 1. Thus, $E[C^2] = E[c^2] = 1$.

By a common reflection argument, BM reaching at least as high as $x > 0$ and ending up at $y = x - (x - y) \in (0, x)$ is tantamount to ending up at $x + (x - y)$. Or, $P\big(H > x,\, C \in [y, y + dy]\big) = P\big(C \in [2x - y,\, 2x - y + dy]\big) = 2\phi\,(2x - y)\,dy$, where $\phi(\cdot) = \frac{1}{\sqrt{2\pi}} \exp\big\{-\frac{1}{2}(\cdot)^2\big\}$ is the standard normal density function (see Siegmund or expression (A.1) in the Appendix for a generalization to (C, H, L)).

Similarly, $P\big(L < z,\, C \in [y, y + dy]\big) = P\big(C \in [2z - y,\, 2z - y + dy]\big) = 2\phi\,(2z - y)\,dy$. Hence, the joint density of H and C is

$$f_{H,C}(x, y) = 4\,(2x - y)\,\phi(2x - y)\,, \qquad 0 < y < x\,, \tag{2.1}$$

and that of L and C is

$$f_{L,C}(z, y) = 4\,(y - 2z)\,\phi(y - 2z)\,, \qquad z < 0 < y\,. \tag{2.2}$$

These joint densities, essentially re-phrasings of a well known formula for the joint density of $(h, h - c)$ (see Yor [12]), lead to the first four of the following five second moments. The fifth is taken from Garman & Klass. Details are omitted. $E[C^2] = 1$ by assumption.

(2.3)

$$E[H^2] = \frac{7}{4}\,, \quad E[L^2] = \frac{1}{4}\,, \quad E[CH] = \frac{5}{4}\,, \quad E[CL] = -\frac{1}{4}\,, \quad E[HL] = 1 - 2\log(2)\,.$$

As a corollary,

Lemma 2.1. *The variance estimators $\hat{\sigma}_i$, $i = 1, 2, 3, 4$ (see (1.1)) are unbiased.*

Seshadri's [10] theorem that $2h(h-c)$ is exponentially distributed with mean 1, and is independent of c, implies that $2H(H-C)$ is exponentially distributed with mean 1, and is independent of C. This is so, simply because the conditional distribution of (h, c) given that $c > 0$ is the (unconditional) distribution of (H, C).

Of course, the same applies to $2l(l-c)$ and $2L(L-C)$. However, $2H(H-C)$ and $2L(L-C)$ are dependent (identities (2.5) yield correlation $1 + \frac{7}{2}\zeta(3) - 8\log(2) = -0.3380$ between the two), and dependent given C.

Otherwise, it would have been very easy to sample (C, H, L) triples. As things stand, it is easy to sample pairs (c, h) (and (c, l)) or (C, H) (and (C, L)), by independently sampling c and $h(h-c)$. A practical approximate method to sample (C, H, L) triples is to sample (C', H') correctly, then make the wrong choice $L' = C' - H'$, not on $[0, 1]$ but on each of the N sub-intervals $[\frac{i-1}{N}, \frac{i}{N}]$. The construction is correct except if H and L are attained in the same sub-interval, the probability of which decreases fast as N increases. Instead of letting $L' = C' - H'$, other copulas may be used, to better approximate features of the joint distribution of (C', H', L').

2.2. The MLE's of σ^2 based on (C, H) and on (C, L) are unbiased

It may be of interest to notice that (2.1) (resp. (2.2)), reinterpreted as $f_{H,C}(x, y; \sigma) = 4\frac{2x-y}{\sigma^3}\phi\left(\frac{2x-y}{\sigma}\right)$, identifies the MLE of σ^2 based on (C, H) (resp. (C, L)) as the average over the sample of $\frac{1}{3}(2H-C)^2 = \frac{1}{3}C^2 + \frac{1}{3}[4(H-C)^2] + \frac{1}{3}[4C(H-C)]$ and $\frac{1}{3}(2L-C)^2 = \frac{1}{3}C^2 + \frac{1}{3}[4L^2] + \frac{1}{3}[-4CL]$. The average of the two, the simple average of the first three unbiased estimators in (1.1), achieves variance 0.3694, above Garman–Klass'.

2.3. The fourth moments of (C, H, L)

The following fourth moments are derived from the joint densities of (H, C) and (L, C). $E[C^4] = 3$ is Gaussian kurtosis.

$$
\begin{aligned}
&E[H^4] = \frac{93}{16}, \quad E[L^4] = \frac{3}{16}, \quad E[CH^3] = \frac{147}{32}, \quad E[CL^3] = -\frac{3}{32}, \\
&E[C^3H] = \frac{27}{8}, \quad E[C^3L] = -\frac{3}{8}, \quad E[C^2H^2] = \frac{31}{8}, \quad E[C^2L^2] = \frac{1}{8}.
\end{aligned}
\tag{2.4}
$$

The following fourth moment information is taken from Garman & Klass. ζ is Riemann's zeta function, with $\zeta(3) = \sum_{k=1}^{\infty} \frac{1}{k^3} \approx 1.2020569$.

$$
\begin{aligned}
E[H^2L^2] &= E[h^2l^2] = 3 - 4\log(2) ,\\
E[C^2HL] &= E[c^2h\,l] = 2 - 2\log(2) - \frac{7}{8}\zeta(3) ,\\
E[H^3L] + E[HL^3] &= E\big[h\,l\,(h^2+l^2)\big] = 6 - 6\log(2) - \frac{9}{4}\zeta(3) ,\\
E[CH^2L] + E[CHL^2] &= E\big[c\,h\,l\,(h+l)\big] = \frac{9}{2} - 4\log(2) - \frac{7}{4}\zeta(3) .
\end{aligned}
\tag{2.5}
$$

There is one more (C,H,L)-based fourth moment needed, whose value does not follow from Garman & Klass'.

Lemma 2.2. $E[CHL^2] = \zeta(3)/16 - 2\log(2) + \frac{47}{32} \approx 0.1575842$.

A proof of Lemma 2.2 can be found in the Appendix. Large sample empirical estimation of $E[CHL^2]$ gave 0.15762, yielding $\mathrm{Var}(\hat{\sigma}_4^2)$ very close to 1. Had $E[CHL^2]$ been equal to $\log(2)(3-4\log(2)) \approx 0.15763$ (initial conjecture), $\mathrm{Var}(\hat{\sigma}_4^2)$ would have been exactly 1.

From all the fourth moments above,

$$
\begin{aligned}
E[C^4] &= 3 ,\\
E[\delta^4] &= E[L^4] = \frac{3}{16} ,\\
E[C\delta^3] &= -E[CL^3] = \frac{3}{32} ,\\
E[C^2\delta^2] &= E[C^2L^2] = \frac{1}{8} ,\\
E[C^3\delta] &= -E[C^3L] = \frac{3}{8} ,\\
E[C^2\Delta\delta] &= E[C^3L] - E[C^2HL] = 2\log(2) + \frac{7}{8}\zeta(3) - \frac{19}{8} ,\\
E[C\Delta\delta^2] &= E[CHL^2] - E[C^2L^2]\\
&= E[CHL^2] - \frac{1}{8} = \frac{\zeta(3)}{16} - 2\log(2) + \frac{43}{32} ,\\
E[\Delta^2\delta^2] &= E[H^2L^2] + E[C^2L^2] - 2E[CHL^2] = \frac{3}{16} - \frac{\zeta(3)}{8} ,\\
2E[\Delta^3\delta] &= E[\Delta^3\delta] + E[\Delta\delta^3]\\
&= -\big(E[H^3L] + E[HL^3]\big)\\
&\quad + E[C^3L] + E[CL^3] - 3E[C^2HL] + 3E[CH^2L] ,\\
&= 6\log(2) - \frac{9}{16}\zeta(3) - \frac{27}{8} .
\end{aligned}
\tag{2.6}
$$

2.4. The covariance matrix of the four basic estimators

Let Σ stand for the covariance matrix of the four basic estimators. Their variances are on the diagonal, their covariances off the diagonal.

Applying the formulas of the previous sub-section, the variances of the basic estimators $\hat{\sigma}_i^2$ (see (1.1)) are

$$\begin{aligned}
\Sigma(1,1) &= \mathrm{Var}(\hat{\sigma}_1^2) = 8\big(E[\delta^4] + E[\Delta^2\delta^2]\big) - 1 = 2 - \zeta(3) = 0.797943\ , \\
\Sigma(2,2) &= \mathrm{Var}(\hat{\sigma}_2^2) = 3 - 1 = 2\ , \\
\Sigma(3,3) &= \mathrm{Var}(\hat{\sigma}_3^2) = 8\big(E[C^2\delta^2] + E[C^2\Delta\delta]\big) - 1 \\
&= 8\left[\log(4) + \frac{7}{8}\zeta(3) - \frac{9}{4}\right] - 1 = 0.504753\ , \\
\Sigma(4,4) &= \mathrm{Var}(\hat{\sigma}_4^2) = \frac{E[\Delta^2\delta^2]}{\big(\log(4) - \frac{5}{4}\big)^2} - 1 = \frac{\frac{3}{16} - \frac{\zeta(3)}{8}}{\big(\log(4) - \frac{5}{4}\big)^2} - 1 = 1.004876\ .
\end{aligned} \tag{2.7}$$

The covariances of the basic estimators are

$$\begin{aligned}
\Sigma(1,2) &= \mathrm{Cov}(\hat{\sigma}_1^2, \hat{\sigma}_2^2) = 4E[C^2\delta^2] - 1 = -\frac{1}{2}\ , \\
\Sigma(1,3) &= \mathrm{Cov}(\hat{\sigma}_1^2, \hat{\sigma}_3^2) = 8E[C\delta^3] + 8E[C\Delta\delta^2] - 1 \\
&= \frac{21 + \zeta(3)}{2} - 16\log(2) = 0.010674\ , \\
\Sigma(1,4) &= \mathrm{Cov}(\hat{\sigma}_1^2, \hat{\sigma}_4^2) = \frac{4E[\Delta\delta^3]}{\log(4) - \frac{5}{4}} - 1 \\
&= \frac{12\log(2) - \frac{27}{4} - \frac{9}{8}\zeta(3)}{\log(4) - \frac{5}{4}} - 1 = 0.580786\ , \\
\Sigma(2,3) &= \mathrm{Cov}(\hat{\sigma}_2^2, \hat{\sigma}_3^2) = 4E[C^3\delta] - 1 = \frac{1}{2}\ , \\
\Sigma(2,4) &= \mathrm{Cov}(\hat{\sigma}_2^2, \hat{\sigma}_4^2) = \frac{E[C^2\Delta\delta]}{\log(4) - \frac{5}{4}} - 1 = \frac{\frac{7}{8}\zeta - \frac{9}{8}}{\log(4) - \frac{5}{4}} = -0.537074\ , \\
\Sigma(3,4) &= \mathrm{Cov}(\hat{\sigma}_3^2, \hat{\sigma}_4^2) = \frac{4E[C\Delta^2\delta]}{\log(4) - \frac{5}{4}} - 1 \\
&= \frac{\frac{\zeta(3)}{4} + \frac{43}{8} - 8\log(2)}{\log(4) - \frac{5}{4}} - 1 = -0.043711\ .
\end{aligned} \tag{2.8}$$

2.5. Derivation of the proposed estimator

Letting α (see (1.3)) stand for the weights assigned to the basic estimators, the weighted sum has variance $\alpha^{\mathrm{T}} \Sigma \alpha$ and mean $\alpha^{\mathrm{T}} \mathbf{1}$. Using a Lagrange multiplier to constrain the mean to be 1, minimal variance is achieved at $\alpha = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^{\mathrm{T}} \Sigma^{-1} \mathbf{1}}$, yielding the weights displayed in (1.3). The variance of the proposed estimator is $\frac{1}{\mathbf{1}^{\mathrm{T}} \Sigma^{-1} \mathbf{1}} = 0.258658$, with corresponding efficiency $2\, \mathbf{1}^{\mathrm{T}} \Sigma^{-1} \mathbf{1} = 7.73221$.

3. HEAVY TAILED RANDOM WALKS — SIMULATION RESULTS

If the logarithmic return process is not distributed as a mean-zero Brownian Motion, variance estimators that are quadratic in S_1 or S_2 can only be compared empirically, aided by simulation. Even the simplest non-Gaussian Lévy process, Poisson process with drift, seems to defy analysis. This section illustrates the empirical construction of quadratic estimators via Regression. We generate power-law-tailed random walk data by assigning quite arbitrarily a t-distribution to its increments. This will permit to monitor comparative performance of the S_1 and S_2 statistics in term of tail thickness.

As is commonly observed in financial data, the logarithmic increments of returns have power-law tails, at least in the visible range, with tail parameter around 3. This means finite variance but infinite variance of the usual empirical variance estimators. Suppose that the basic process on which (Open, Close, Min, Max) data is reported per time window is a random walk with t-distributed increments. A simulation analysis will now be reported, in which the number of increments of the random walk per time window is $10, 30$ and 50, and the degrees of freedom (df) range from 1.5 to 5 with step size 0.5. Minimum sum-of-squares quadratic functions with mean 1 of the S_1 and S_2 statistics were fitted by Regression, with sample size 10^5: the regression coefficients were identically calibrated so that the predictor of unity has mean 1 in each such sample. Each such Regression was repeated 100 times, and the averages of the corresponding regression coefficients and overall "variances" were recorded. Of course, second moments are finite only for $df > 2$ and fourth moments are finite only for $df > 4$, but the empirical study seems instructive. A sample of size 10^5 from the sum of $N = 50$ $t_{\{df=3\}}$-distributed random variables typically displays lighter tails than $df = 3$ would entail. Table 1 reports the empirical minimum variance of the quadratic functions, and Table 2 reports the coefficients of the building blocks of expression (1.1) that yield the minimum-variance quadratic function for each case. These building blocks have expectation 1 for Brownian Motion but not for random walk, so their coefficients need not add up to unity. Table 1 displays performances similar to those derived for Brownian Motion for moderate df, fast deteriorating when *df* decreases, in which case S_2 data progressively outperforms S_1 data. S_2 data yields lower variances than S_1 data throughout the range, as well as for uniform and double exponentially distributed increments, although the difference in variance in these light-tail cases is as small as for *BM*.

Table 1: Minimum variance of mean-1 quadratic functions of S_1 and S_2 data.

df	$N=10$		$N=30$		$N=50$	
	S_2	S_1	S_2	S_1	S_2	S_1
1.5	16.2403	51.0366	8.3438	32.4697	6.5322	28.3950
2.0	4.8444	6.6039	2.6532	3.8327	2.1972	3.2252
2.5	2.5864	2.8365	1.4297	1.5529	1.1718	1.2627
3.0	1.7359	1.8038	0.9527	0.9782	0.7630	0.7788
3.5	1.2334	1.2746	0.6809	0.6991	0.5467	0.5624
4.0	0.9469	0.9776	0.5409	0.5585	0.4532	0.4686
4.5	0.7864	0.8124	0.4792	0.4957	0.4094	0.4239
5.0	0.7071	0.7296	0.4473	0.4629	0.3896	0.4037
∞	0.4679	0.4826	0.3630	0.3765	0.3369	0.3496
∞, $N=\infty$					0.2587	0.27

It is of interest to observe how does S_2 outperform S_1 data for low *df*. Table 2 shows that the role of C is downplayed or even dampened in favor of those of $H-C$ and $-L$, gradually incorporating C into the Regression as *df* increases. The rationale for this is that the tail parameter of sums of i.i.d. data is the same as that of the summands, whereas the tail parameter of extrema is the sum of those of the summands.

Table 2: Coefficients of the minimum variance mean-1 quadratic function of S_2 data for $N=50$ increments per time window.

df	$N=50$			
	$2\big((H-C)^2+L^2\big)$	C^2	$2\,(H-C-L)\,C$	$\dfrac{-(H-C)L}{2\log(2)-5/4}$
1.5	0.0209	−0.0000	0.0010	0.1724
2.0	0.1358	−0.0004	0.0352	0.1561
2.5	0.1745	−0.0034	0.1573	0.1215
3.0	0.1827	0.0140	0.2461	0.1149
3.5	0.2006	0.0666	0.2460	0.1228
4.0	0.2185	0.1081	0.2442	0.1317
4.5	0.2335	0.1271	0.2620	0.1399
5.0	0.2480	0.1395	0.2781	0.1473
∞	0.3974	0.2321	0.4390	0.2245
∞, $N=\infty$	0.2736	0.1604	0.3652	0.2009

This makes C theoretically as heavy tailed as each increment, but makes $H-C$ and $-L$ have lighter tails than the increments. In contrast, the $[h,c,l]$ data of statistic S_1 is less able to split variables into light tail and heavy tail components. Although $h-|c|-l = H-C-L$, the insistence on resorting to quadratic functions leaves it out of the S_1 game. Still, both statistics seem to work fairly well even under low df. In contrast to the variances 2.1972 or 3.2252 for $df=2$, 0.7630 or 0.7788 for $df=3$ and 0.4532 or 0.4686 for $df=4$ (see $N=50$ in Table 1), the calibrated C^2 has respective empirical variance above 5000, 16 and 2.5, converging reasonably fast $\bigl(2+\frac{6}{(df-4)N}\bigr)$ to 2 thereafter.

APPENDIX — PROOF OF LEMMA 2.2

For the sake of conciseness, the tedious integration to be presented will be restricted to the identification of $E[CHL^2]$, although, in principle, more general joint moments and moment generating function of (C,H,L) could have been identified.

Consider the infinitesimal event $\bigl\{BM(1)\in(\xi,\xi+d\xi),\ BM(s)\in(a,b),\ \forall s\in[0,1]\bigr\}$, where $a<\min(\xi,0)\le 0\le\max(\xi,0)<b$. By Siegmund's Corollary 3.43, its probability $Q(\xi,a,b)\,d\xi$ is as follows

$$Q(\xi,a,b) \;=\; \sum_{j=-\infty}^{\infty}\Bigl\{\phi\bigl(\xi-2j\,(b-a)\bigr)-\phi\bigl(\xi-2a-2j\,(b-a)\bigr)\Bigr\}\,. \tag{A.1}$$

The joint density $f_{c,h,l}(\xi,a,b)$ is (minus) the mixed second derivative of Q with respect to a and b, on $\bigl\{\xi\in(a,b),\ a<0,\ b>0\bigr\}$. The joint density $f_{C,H,L}$ is simply $2\,f_{c,h,l}$, restricted to $\bigl\{\xi\in(0,b),\ a<0,\ b>0\bigr\}$. The two terms in the $j=0$ and second term in the $j=1$ summands vanish because they are independent of at least one of a and b.

To calculate $E[CHL^2]$, the contribution of each summand in (A.1) will be integrated in three univariate steps. The first step will integrate over $a\in(-\infty,0)$ the product of a^2 and the pertinent mixed second derivative. $\frac{\partial}{\partial a}\phi\,(\xi+Ka+Mb)\,da$ is to be interpreted as the integration-by-parts element $d\phi\,(\xi+Ka+Mb)$, viewed as a function of a.

$$\begin{aligned}
&\int_{-\infty}^{0}\frac{\partial}{\partial b}a^2\,\frac{\partial}{\partial a}\phi\,(\xi+Ka+Mb)\;da \;=\\
&\qquad=\;\frac{2}{K^2}\,\frac{\partial}{\partial b}\Bigl[\phi(\xi+Mb)+(\xi+Mb)\,\Phi(\xi+Mb)\Bigr]\quad(\text{for } K>0)\\
&\qquad=\;\frac{2M}{K^2}\;\Phi(\xi+Mb)\qquad\qquad(\text{for } K>0)\\
&\qquad=\;\frac{2M}{K^2}\;\Phi(\xi+Mb)-\frac{2M}{K^2}\quad(\text{for } K<0)\,.
\end{aligned}\tag{A.2}$$

Now expression (A.2) will be multiplied by ξ and integrated over $\xi \in (0, b)$. For $K > 0$ ($K < 0$) it is convenient to integrate Φ^* (Φ). These terms appear in (A.3) and (A.4). The free term in (A.2) contributes $\frac{2M}{K^2}\frac{b^2}{2}$ and cancels with the corresponding b^2 term in (A.4).

$$
\begin{aligned}
&\int_0^b \xi \frac{\partial}{\partial b} \int_{-\infty}^0 a^2 \phi(\xi + K\,da + Mb)\; d\xi \;= \\
&\qquad = \frac{2M}{K^2} \int_{Mb}^{(M+1)\,b} y\, \Phi(y)\; dy \;-\; \frac{2M^2 b}{K^2} \int_{Mb}^{(M+1)\,b} \Phi(y)\; dy \\
&\qquad = \frac{M}{K^2} \Big[(M^2 b^2 + 1)\, \Phi(Mb) - \big((M^2-1) b^2 + 1\big)\, \Phi\big((M+1)\,b\big) \\
&\qquad\quad + Mb\, \phi(Mb) - (M-1)\, b\, \phi\big((M+1)\,b\big)\Big] \\
&\text{(A.3)} \qquad = -\frac{M}{K^2} \Big[(M^2 b^2 + 1)\, \Phi^*(Mb) - \big((M^2-1) b^2 + 1\big)\, \Phi^*\big((M+1)\,b\big) \\
&\text{(A.4)} \qquad\quad + Mb\, \phi(Mb) - (M-1)\, b\, \phi\big((M+1)\,b\big)\Big] + \frac{M}{K^2} b^2\,.
\end{aligned}
$$

Finally, expressions (A.3) and (A.4), multiplied by b and integrated over $b \in (0, \infty)$, via

$$
\text{(A.5)} \quad \int_0^\infty b^3\, \Phi^*(Ab)\, db = \frac{3}{8A^4}\,, \quad \int_0^\infty b\, \Phi^*(Ab)\, db = \frac{1}{4A^2}\,, \quad \int_0^\infty b^2\, \phi(Ab)\, db = \frac{1}{2A^3}\,,
$$

yield a rational function of j (with $M = 2j$ and $K = -2j$ or $K = -2(j-1)$) whose sum contains only terms of the form $-\sum_1^\infty (-1)^j \frac{1}{j} = \log(2)$ and $\sum_1^\infty \frac{1}{j^3} = \zeta(3)$, as in the statement of Lemma 2.2. Further details are omitted.

ACKNOWLEDGMENTS

The topic under study was motivated by a project at ISTRA Research, Israel. The collaboration of Shlomo Ahal, Jonathan Lewin and Alon Wasserman is greatly appreciated. Ahal's careful reading and constructive comments are an essential part of the paper. Warm thanks are extended to my hosts at Columbia University's Statistics Department during a sabbatical visit in the Spring of 2008.

REFERENCES

[1] Alizadeh, S.; Brandt, M. and Diebold, F.X. (1980). Range-based estimation of stochastic volatility models, *Journal of Finance*, **57**, 1047–1091.

[2] Ball, C.A. and Torous, W.N. (1984). The Maximum Likelihood estimation of security price volatility: theory, evidence and application to option pricing, *The Journal of Business*, **57**, 97–112.

[3] Blackwell, D. (1947). Conditional expectation and unbiased sequential estimation, *The Annals of Mathematical Statistics*, **18**, 105–110.

[4] Christensen, K.; Podolskij, M. and Vetter, M. (2009). Bias-correcting the realized range-based variance in the presence of market microstructure noise, *Finance and Stochastics*, **13**, 239–268.

[5] Garman, M.B. and Klass, M.J. (1980). On the estimation of security price volatilities from historical data, *The Journal of Business*, **53**, 67–78.

[6] Joshi, V.M. (1976). On the attainment of the Cramér–Rao lower bound, *The Annals of Statistics*, **4**, 998–1002.

[7] Magdon-Ismail, M. and Atiya, A.F. (2001). A maximum likelihood approach to volatility estimation for a Brownian motion using the high, low and close, *Quantitative Finance*, **1**, 1–9.

[8] Rao, C.R. (1946). Minimum variance and the estimation of several parameters, *Proceedings of the Cambridge Philosophical Society*, **43**, 280–283.

[9] Rogers, L.C.G. and Satchell, S.E. (1991). Estimating variance from high, low and closing prices, *Annals of Applied Probability*, **1**, 504–512.

[10] Seshadri, V. (1988). Exponential models, Brownian Motion and independence, *Canadian Journal of Statistics*, **16**, 209–221.

[11] Siegmund, D.O. (1985). *Sequential Analysis: Tests And Confidence Intervals*, Springer Series in Statistics, Springer Verlag, New York.

[12] Yor, M. (1997). Some remarks about the joint law of Brownian Motion and its supremum, *Séminaire de Probabilités (Strasbourg)*, **31**, 306–314.